Clim Dyn (2007) 29:231–250
DOI 10.1007/s00382-007-0232-2

GCM intercomparison of global cloud regimes: present-day evaluation and climate change response

K. D. Williams · G. Tselioudis

Received: 15 June 2006 / Accepted: 20 January 2007 / Published online: 2 March 2007

Abstract The radiative feedback from clouds remains the largest source of variation in climate sensitivity amongst general circulation models (GCMs). A cloud clustering methodology is applied to six contemporary GCMs in order to provide a detailed intercomparison and evaluation of the simulated cloud regimes. By analysing GCMs in the context of cloud regimes, processes related to particular cloud types are more likely to be evaluated. In this paper, the mean properties of the global cloud regimes are evaluated, and the cloud response to climate change is analysed in the cloud-regime framework. Most of the GCMs are able to simulate the principal cloud regimes, however none of the models analysed have a good representation of trade cumulus in the tropics. The models also share a difficulty in simulating those regimes with cloud tops at mid-levels, with only ECHAM5 producing a regime of tropical cumulus congestus. Optically thick, high top cloud in the extra-tropics, typically associated with the passage of frontal systems, is simulated considerably too frequently in the ECHAM5 model. This appears to be a result of the cloud type persisting in the model after the meteorological conditions associated with frontal systems have ceased. The simulation of stratocumulus in the MIROC GCMs is too extensive, resulting in the tropics being too reflective. Most of the global-mean cloud response to doubled CO_2 in the GCMs is found to be a result of changes in the cloud radiative properties of the regimes, rather than changes in the relative frequency of occurrence (RFO) of the regimes. Most of the variance in the global cloud response between the GCMs arises from differences in the radiative response of frontal cloud in the extra-tropics and from stratocumulus cloud in the tropics. This variance is largely the result of excessively high RFOs of specific regimes in particular GCMs. It is shown here that evaluation and subsequent improvement in the simulation of the present-day regime properties has the potential to reduce the variance of the global cloud response, and hence climate sensitivity, amongst GCMs. For the ensemble of models considered in this study, the use of observations of the mean present-day cloud regimes suggests a potential reduction in the range of climate sensitivity of almost a third.

Electronic supplementary material The online version of this article (doi:10.1007/s00382-007-0232-2) contains supplementary material, which is available to authorized users.

K. D. Williams (✉)
Hadley Centre, Met Office,
FitzRoy Road, Exeter EX1 3PB, UK
e-mail: keith.williams@metoffice.gov.uk

K. D. Williams
Environmental Systems Science Centre,
University of Reading, FitzRoy Road,
Exeter EX1 3PB, UK

G. Tselioudis
NASA Goddard Institute for Space Studies,
New York, NY, USA

G. Tselioudis
Department of Applied Physics,
Columbia University, New York, NY, USA

1 Introduction

General circulation models (GCMs) are the primary tools used for climate change prediction. However, leading-order measures of the climate change response, such as climate sensitivity, vary between state-of-the-art GCMs (e.g.

Cubasch et al. 2001). It is desirable to reduce this uncertainty if policy makers are to have confidence in making potentially costly socio-economic decisions. Comparison studies suggest that differences in the radiative feedback from clouds account for much of the variation in climate sensitivity between current GCMs (Cess et al. 1990; Senior and Mitchell 1993; Webb et al. 2006; Soden and Held 2006; Ringer et al. 2006).

Evaluation of cloud in GCMs traditionally involved comparing climatological maps of cloud variables (typically cloud radiative forcing, CRF, or total cloud amount) simulated by the GCM, with observational data. Whilst it is desirable that a model should be able to accurately simulate the mean present-day climate, this does not necessarily imply that the model will correctly simulate feedback processes relevant to climate change. Recently, evaluation has become more aligned with atmospheric processes, with techniques being developed which aim to separate clouds into particular 'regimes' before evaluating them. Various types of regime have been studied including dynamical regimes based on vertical velocity (e.g. Bony et al. 2004; Norris and Weaver 2001), surface pressure regimes (Tselioudis et al. 2000), and physico-dynamical regimes based on a combination of vertical velocity with SST and/or lower tropospheric stability (e.g. Ringer and Allan 2004; Williams et al. 2003, 2006). Jakob and Tselioudis (2003) apply a statistical clustering technique to International Satellite Cloud Climatology Project (ISCCP) data over the tropical warm pool, to separate clouds into regimes with similar cloud-top pressures (CTP), cloud optical depths (τ) and total cloud covers (TCC). Rossow et al. (2005) extend the analysis to provide a detailed examination of the observed cloud regimes throughout the tropics, whilst Williams et al. (2005) apply the clustering technique to two GCMs and investigate the climate change response over four specific geographical regions. This study builds on that of Rossow et al. (2005) and Williams et al. (2005) by applying the clustering methodology globally to six GCMs, for which daily ISCCP simulator data (Klein and Jakob 1999; Webb et al. 2001) have been submitted to the Cloud Feedback Model Intercomparison Project (CFMIP; McAvaney and Le Treut 2003; http://www.cfmip.net). The present-day characteristics of the cloud regimes are evaluated and the climate change response is analysed.

The models, observational data and clustering technique are described in the next section. The characteristics of the mean present-day cloud regimes, as simulated by the GCMs, are evaluated in Sect. 3. In Sect. 4 the climate change response of the different GCMs are compared. The potential of the regime evaluation to constrain the range of climate sensitivity from the GCMs is investigated in Sect. 5. A summary and discussion are in Sect. 6.

2 Models, observational data and methodology

2.1 Model description and experimental design

This study uses equilibrium control and $2 \times CO_2$ atmosphere—mixed-layer ocean (slab model) experiments from GCMs which have been submitted to CFMIP. An essential requirement of CFMIP is that models submitted include the ISCCP simulator (Klein and Jakob 1999; Webb et al. 2001; http://www.gcss-dime.gis.nasa.gov/simulator.html). This code is designed to be embedded in a GCM's radiation scheme to produce diagnostics which can be compared directly with the ISCCP observational products. To date, the daily ISCCP simulator diagnostics and top of atmosphere (TOA) fluxes which are required for this study have been submitted for six atmosphere–slab ocean models. These are listed in Table 1 and together form an ensemble of GCMs with some significant structural differences, covering a factor of 2 in horizontal and vertical resolution. Although three Hadley Centre models are included, HadGSM1 and HadSM3 (slab-model versions of HadGEM1 and HadCM3) have considerable structural differences, including a different dynamical core, resolution and many different or revised physical parametrizations. HadSM4 is an intermediate model, containing aspects of both HadSM3 and HadGSM1. The lower and higher sensitivity versions of the MIROC 3.2 model differ in two respects: the temperature range over which mixed phase cloud is simulated, and the fact that melted cloud ice is rained-out in MIROC-lo whilst it is converted to cloud water in MIROC-hi (Ogura et al. 2007). The MIROC-lo model is the same as the 'medium resolution' MIROC model submitted to the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report database. MIROC-hi is at the same resolution as MIROC-lo with the 'hi' referring to sensitivity. It is *not* the same model as the high resolution version of the MIROC model which has been submitted to the IPCC Fourth Assessment Report database.

The surface temperature of each of the slab-ocean models is maintained close to climatological values, in the absence of ocean currents, by use of a monthly varying heat flux. This is calculated in a calibration experiment (performed for each model prior to the main control and $2 \times CO_2$ experiments) in which the SSTs are reset to climatological values at each timestep. Five years of daily data from the control and $2 \times CO_2$ simulations at equilibrium are analysed in this study. An exception is for the MIROC models, for which only 2 years of all the necessary daily diagnostics have been submitted. Analysis of the other models indicates that the results change very little between using 2 and 5 years of daily data.

Table 1 List of models used in this study (all are atmosphere—mixed-layer ocean configurations of the model)

Model	Resolution	Nature of cloud scheme	Main references
ECHAM5	T63 L32	Prognostic	Roeckner et al. (2003)
HadSM3	N48 L19	Diagnostic	Pope et al. (2002); Williams et al (2001)
HadSM4	N48 L38	Ddiagnostic	Webb et al. (2001)
HadGSM1	N96 L38	Diagnostic	Martin et al. (2006); Johns et al (2006)
MIROC-lo	T42 L20	Diagnostic	K-1 model developers 2004
MIROC-hi	T42 L20	Diagnostic	K-1 model developers 2004

Horizontal resolution is pre-fixed by '*T*' for the truncation of spectral models and '*N*' for half the number of east–west points for grid-point models (this notation permits approximate comparison of the two model structures). The number of atmosphere levels is prefixed by '*L*'. Also shown is whether the model uses prognostic or diagnostic cloud liquid water in its large-scale cloud scheme

2.2 Observational data

The cloud clusters produced from the GCMs are compared with ISCCP observational data (Rossow and Schiffer 1991). The observational clusters are produced using the ISCCP D1 product which contains cloud amount in six τ and seven CTP categories every 3 h and are re-gridded onto a 2.5° grid (i.e. the dataset is formed of a τ-CTP histogram for each grid-point). The data are averaged to daily means for comparison with the daily data available from the models. Similar results are obtained when using 3-hourly or daily data. The ISCCP D1 product and ISCCP simulator diagnostics are only available at sun-lit points, so the day-mean is the average of the 3-hourly samples during daylight hours. Therefore, a significant bias in the diurnal cycle of a model may bias the evaluation.

Cloud tops in ISCCP are derived from the thermal infrared radiance of a cloudy pixel, which is translated to a pressure using the TIROS Operational Vertical Sounder (TOVS) temperature and humidity profiles. The presence of an optically thin cloud above an optically thick one would lift the pixel infrared emission level to the middle troposphere resulting in the retrieval of a middle level cloud in the place of the two overlapping cloud layers. The use of the ISCCP simulator on the model cloud parameters should account for this effect as it also outputs a middle level cloud in a similar situation. In addition to this effect, any systematic biases in the TOVS temperature and humidity profiles would result to systematic errors in the ISCCP cloud top pressures. Wang et al. (1999) show that such TOVS errors produce cloud tops that are 50–60 hPa too high in the Atlantic stratocumulus region.

Two datasets are used for derived observations of CRF: the ISCCP FD product (Zhang et al. 2004) and the S4G product from the Earth Radiation Budget Experiment (ERBE). In the case of ERBE, clear-sky measurements are only available on each day at places which are cloud-free, hence monthly-mean clear-sky fluxes were subtracted from the daily all-sky fluxes in order to produce a daily CRF. It should be noted that unlike ISCCP FD, the ERBE dataset contains a reasonable amount missing data, particularly over land, which may lead to some bias in the CRF of regimes in these areas.

Five years of observational data for the period March 1985–February 1990 are used for evaluation of the GCMs. This period includes a large El-Niño southern oscillation (ENSO) cycle. Since slab models do not simulate the coupled atmosphere–ocean processes of the ENSO, the observations are likely to contain more inter-annual variability than the GCMs.

2.3 Clustering methodology

The ISCCP cloud clustering uses the KMEANS clustering algorithm (Anderberg 1973), as described by Jakob and Tselioudis (2003). Each cluster centroid is initially seeded with a random histogram from the full spatio-temporal dataset. The method assigns each spatio-temporal data point to the cluster whose centroid is closest (as measured by the Euclidean distance) to the τ-CTP histogram, then recomputes the cluster centroid. The method works iteratively in order to produce the optimum set of resulting cluster centroids. Thirty iterations have been found to be sufficient to ensure convergence. Unlike Jakob and Tselioudis (2003) and Rossow et al. (2005), the relatively small number of grid-points with completely clear-skies are removed before clustering. This has very little effect on the ISCCP observational data (which has a clear-sky RFO of less than 0.5%), but helps with the stability of the GCM clusters for which the RFO of clear-sky can be up to 14%.

An empirical method to choose the number of clusters is employed here which is one of the criteria used by Rossow et al. (2005) in defining their number of clusters. The method involves repeating the clustering, starting with two clusters, increasing the number (n) of requested clusters by one on each occasion. The process is stopped once the correlation between any two resulting cluster centroid histograms exceeds 0.9 (indicating that a cluster is being repeated—Rossow et al. 2005). The n–1 cluster centroid set is then retained. An additional criterion is employed

that the relative frequency of occurrence (RFO) of the least common regime should not be less than 3.5%, as in these cases the results were found to become sensitive to the initial seeding of the clusters. Clusters produced in this way also satisfy most of the tests for cluster uniqueness proposed by Rossow et al. (2005).

In order to check for sensitivity to the seeding, the clustering was repeated 40 times (each occasion including the 30 iterations to produce the optimum centroids). Identical cluster centroids had to be produced on at least 75% of occasions (100% was found to be too strict as, particularly for the GCMs, some of the datasets appear to contain several cloud regimes which occur very infrequently, but the major regimes tend to be repeated before these very infrequent regimes are separated out).

The clustering of the ISCCP observational data has been carried out using this method for the region 15°N–15°S. The resulting clusters are extremely similar to those obtained by Rossow et al. (2005) (with RFOs within 1%), suggesting that some of the subjective differences in methodology, use of daily rather than 3-hourly data, and using 5 years rather than 21.5 years of data, have little impact on the results.

The TCC and RFO for the resulting clusters is calculated, together with the average shortwave and longwave CRF components (SCRF and LCRF) for spatio-temporal grid points included in each cluster. Uncertainty estimates have been calculated for these quantities using a bootstrap method. The spatio-temporal data-points have been randomly re-sampled 20 times to provide samples of the same size as the original dataset, and the clustering repeated on each. Two standard deviations of the results have been calculated to provide approximately a 95% confidence interval. One degree of freedom has been assumed as the same data are being re-sampled each time, however this assumption is conservative and will tend to slightly overestimate the uncertainty.

The GCM ISCCP simulator data from the control and $2 \times CO_2$ runs are clustered using exactly the same method as for observational data except that the histogram also contains cloud amount in each CTP bin for optical depths which are too small to be detected by ISCCP ($\tau < 0.3$). Therefore the clustering algorithm uses 49 element histograms from the GCMs compared with 42 element histograms for the observational data. The model data are clustered independently of the observational data, hence different numbers of clusters are produced from each model and from the observations (Table 2). Further research is required to understand why some of the clusters are not simulated individually by the GCMs and to correct this deficiency. However, for the present study, to aid comparison between models, several of the clusters are subjectively combined to form a common set of 'principal cloud regimes'. Different numbers of clusters may be produced for the $2 \times CO_2$ simulation as the control, however this was found to be rare and the same principal cloud regimes as identified from the control could be formed in all cases.

Table 2 The number of clusters (excluding 'clear-sky') produced for each region

	Number of tropical clusters	Number of extra-tropical clusters	Number of ice/snow clusters
ISCCP	7	7	6
ECHAM5	5	9	7
HadSM3	4	6	3
HadSM4	6	7	5
HadGSM1	7	8	7
MIROC-lo	3	8	8
MIROC-hi	3	7	9

The clustering and subsequent analysis has been carried out on each model at the resolution of the model. This is to ensure that individual cloud regimes are not combined, particularly for those models where interpolation onto a common grid would not involve an integer number of grid-boxes. In order to test the sensitivity of the clusters to the resolution, the horizontal resolution from HadGSM1 (the highest resolution model used here) was degraded by half (to be equal to the lowest resolution model), and the clustering repeated. The clusters produced were indistinguishable from those produced at the full resolution and the RFO, TCC and CRF were all well within the uncertainty estimate due to variability.

To distinguish between clouds produced by different processes and to isolate observational errors, it has been found necessary to split the global analysis into three geographical regions: the tropics (which is defined here as 20°N–20°S); the ice-free extra-tropics (regions polewards of 20°N/S which are snow/ice-free in the present-day); and regions which are ice/snow-covered in the present-day. The area of the τ-CTP histogram occupied by deep convective cloud is similar to that occupied by frontal cloud (particularly for some of the GCMs), but the processes associated with these two cloud types differ. Hence it is useful to separately cluster the regions which contain each of these cloud types. Twenty degree N/S is found to be the most appropriate geographical separation for these two cloud types, forming tropical and extra-tropical regions. It is also desirable to separate those extra-tropical regions which are ice/snow-covered from those which are ice-free, since the observations are less reliable over high albedo surfaces. In addition, changes in ice/snow cover under climate change may result in a change in CRF. The boundary between the ice-covered and ice-free extra-tropics is allowed to vary

spatio-temporally. Points which are snow or ice-covered for the observational data are identified using the ISCCP-IS product (Rossow et al. 1996). For those GCMs which have submitted sea-ice and snow fractions or amounts, the boundary is determined using the model fields, otherwise (in the case of the MIROC models) the ISCCP-IS product is used.

3 Evaluation of mean present-day cloud regimes

Comparison of the zonal-mean multi-annual-mean SCRF and LCRF from the GCMs with two observational datasets reveals that, in general, the models capture the main features of the latitudinal variation in CRF (Fig. 1). No single model may be considered particularly superior or deficient, however a number of systematic errors can be identified. These include the SCRF in the MIROC models (particularly MIROC-lo) being stronger than observed in the tropics; the mid-latitude SCRF maximum being too strong in ECHAM5 in both hemispheres, and in the MIROC models in the Southern Hemisphere; the LCRF being too weak in HadSM3. This section uses the cloud regime based evaluation in order to identify the cause of some of these systematic errors, and determine whether the overall similarity of the zonal mean CRF is a result of the GCMs simulating cloud regimes with comparable radiative properties, or whether a compensation of different regime characteristics exists between the models.

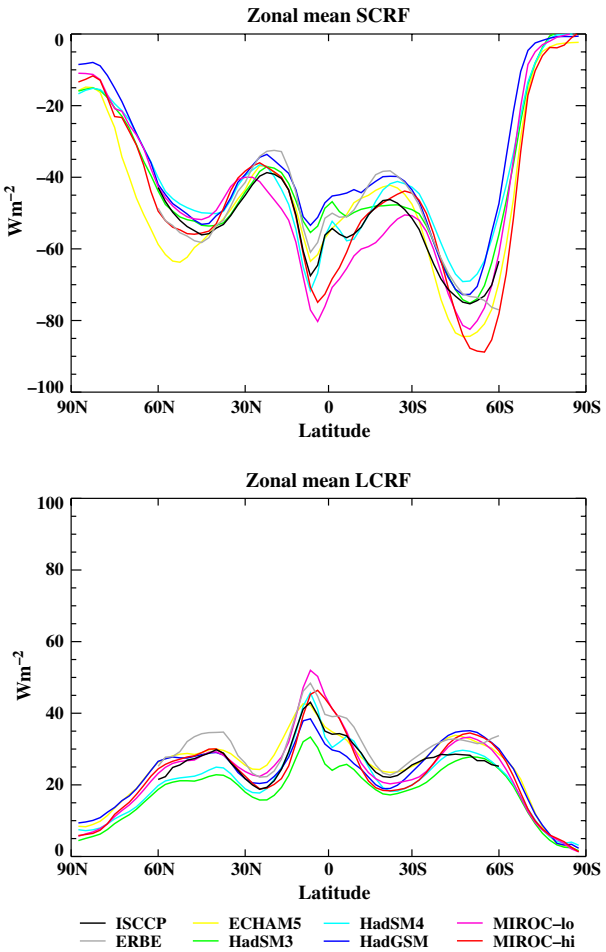

Fig. 1 Zonal-mean multi-annual-mean SCRF and LCRF for each of the GCMs and for observational data from ISCCP FD and ERBE. The observational zonal means are only shown 60°N–60°S due to the datasets being considered less reliable over regions covered by ice/snow

3.1 Tropics

The observed mean histograms for each of the principal cloud regimes in the tropics, and their geographical frequency, are shown in Fig. 2. The regimes have been labelled as 'shallow cumulus'; 'mid-level convective cloud'; 'stratocumulus'; 'deep convective cloud and anvil'; 'thin cirrus'. The regime names indicate the main cloud type which might be expected to occur, although it is recognised that a morphological cloud type cannot be explicitly identified from τ-CTP-TCC alone. Similar histograms and frequency maps for each of the models in each region are available as electronic supplementary material (ESM Figs. 1–8). Here, the main features of the simulated cloud regimes will be summarised.

A regime with a low TCC ('Shallow Cumulus') is simulated by the Hadley Centre models, but not by the other GCMs. ISCCP observations suggest that this regime should be dominated by shallow cumulus in the presence of a small amount of cirrus. Whilst the Hadley Centre models produce some cirrus, there is little evidence of shallow cumulus and Zhang et al. (2005) suggest that this underestimation of shallow cumulus is a general issue for GCMs. Several cumulus clouds may exist within a single satellite pixel with clear-sky between, so ISCCP may underestimate the optical depth and overestimate the TCC of shallow cumulus. If it is assumed that 40% of a pixel is covered by cumulus with clear-sky between (Klein et al. 1995), then the observed ISCCP regime with a maxima of 5.5% cloud cover in the optical depth range 1.3–3.6 should be simulated by the GCMs as being 2.2% cloud cover with an optical depth in the range 5.5–25. However, all of the models which simulate the regime have less than 1% cloud cover in this range, hence satellite beam-filling only accounts for some of the discrepancy.

ECHAM5 is the only GCM analysed which produces a separate regime of cloud with tops at a mixture of different levels, including mid-levels. However, the cloud simulated by the model for this regime is considerably optically thicker than that observed by ISCCP. The lack of an accurate simulation of these two cloud regimes is of con-

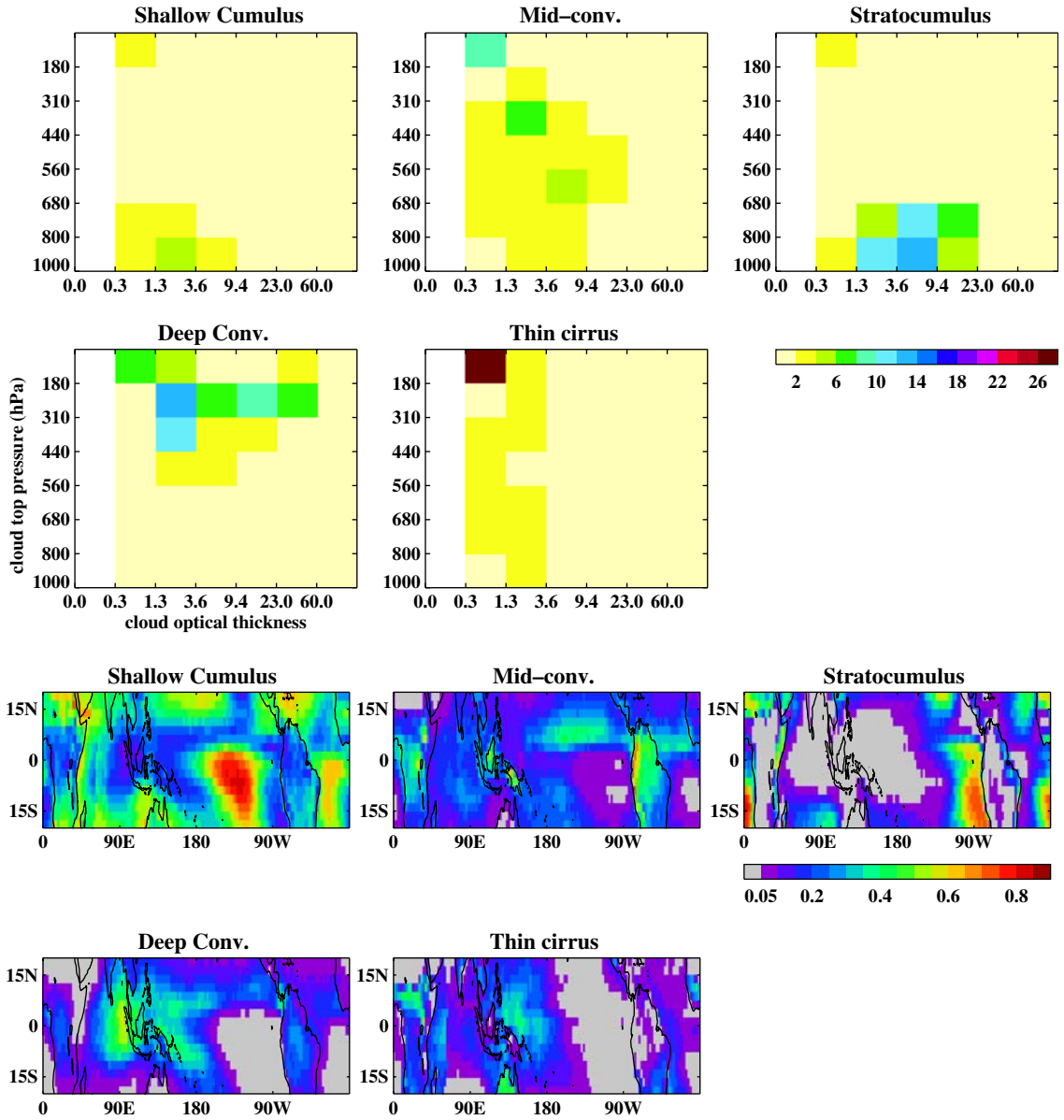

Fig. 2 *Top*: mean CTP-τ regime histograms resulting from clustering daily ISCCP data over the tropics (20°N–20°S). Some of the clusters have been combined to form a set of principal cloud regimes. *Colours* indicate the cloud amount (%) in each CTP-τ category. Note that the *pale yellow* includes cloud amounts up to 2% (including 0%). *Bottom*: the temporal proportion of each grid-box (as a fraction of the total number of days) included in each regime

cern since they may produce a cloud feedback process which is missing from the GCM ensemble.

All of the GCMs except ECHAM5 simulate a stratocumulus regime (Fig. 3), although the properties of the regime show considerable variation amongst the models (Fig. 4). The regime τ-CTP histogram from HadSM4 has the smallest Euclidean distance to the ISCCP observed histogram, which indicates that the simulation of the regime from this model is closest to ISCCP (Table 3). (The Euclidean distance between the simulated and observed histograms for the other regimes are available as ESM Tables 1, 2). In the MIROC models and HadSM3, the stratocumulus is too optically thick, which in the case of HadSM3 results in a SCRF which is around double what is observed (Fig. 4). Both of the CRF components are accurately simulated by HadSM4 whereas most of the other models have an excessive SCRF for this regime. The tropical mean RFO and geographical location of the regime is particularly well simulated in HadSM4. Whilst HadGSM1 simulates stratocumulus with a reasonably similar RFO in the traditional 'stratocumulus regions', the RFO is lower than observed elsewhere in the subtropics. The regime is generally simulated too infrequently in HadSM3 and too often in the MIROC models, partly compensating

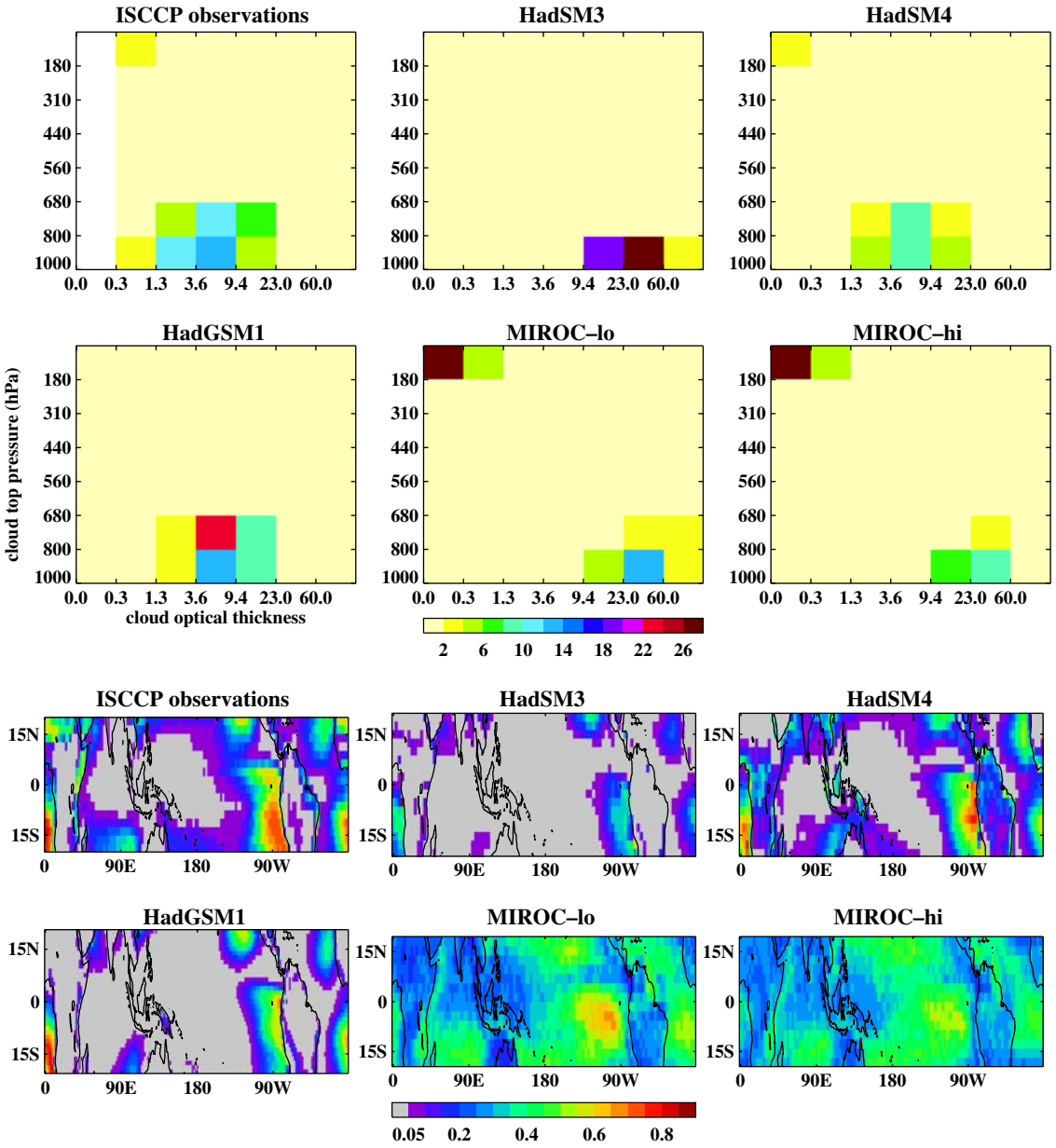

Fig. 3 *Top*: mean CTP-τ histograms for the tropical stratocumulus regime resulting from ISCCP and from each GCM. Bottom: The temporal proportion of each grid-box (as a fraction of the total number of days) included in the tropical stratocumulus regime. (ECHAM5 does not simulate this regime)

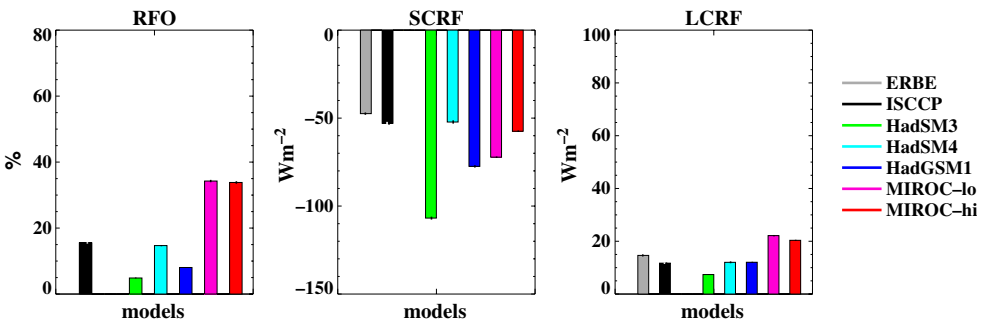

Fig. 4 Mean RFO, SCRF and LCRF for the tropical stratocumulus regime. The small line at the end of each bar represents uncertainty due to inter-annual variability (it does not include uncertainty in the observational dataset)

Table 3 Euclidean distance (%) between the mean τ-CTP regime histograms from each model and those observed by ISCCP for the tropical stratocumulus regime. (ECHAM5 does not simulate the regime

HadSM3	36.6
HadSM4	10.7
HadGSM1	17.7
MIROC-lo	24.2
MIROC-hi	22.7

for their lack of a shallow cumulus regime. This excessive RFO of stratocumulus is the main cause of the tropical SCRF being too strong in the MIROC models (Fig. 1).

Following Jakob et al. (2005), it is useful to examine the meteorological characteristics associated with the regime to aid understanding of the differences in the cloud simulation between the GCMs. Mean thermodynamic profiles for each of the models and two re-analyses are shown in Fig. 5 for the stratocumulus regime. The re-analyses are the European Centre for Medium Range Weather Forecasts (ECMWF) 40 year re-analysis (ERA-40) (http://www.ecmwf.int/products/data/archive/descriptions/e4) and the National Center for Atmospheric Research (NCAR)/National Centers for Environmental Prediction (NCEP) re-analysis (Kalnay et al. 1996). Those data-points assigned to the stratocumulus regime in clustering the ISCCP data are extracted from the re-analyses and averaged. The importance of lower tropospheric stability on the formation of stratocumulus has been highlighted by several studies (e.g. Klein and Hartmann 1993; Williams et al. 2006). The re-analyses and the Hadley Centre models have evidence of a θ_{es} (saturated equivalent potential temperature) inversion at the top of the boundary layer, which distinguishes the regime from the shallow cumulus case (not shown). Such an inversion is present but much weaker in the mean profile for MIROC models, suggesting that the large amount of 'stratocumulus' in these models is not due to a strong inversion over an excessively large area. However, the MIROC boundary layer profiles are the most humid of the GCMs which is likely account for the large amount of low cloud. The temperature and dew point profiles in the lower troposphere for stratocumulus appear closest to the re-analyses in HadSM4, which may be associated with this model having a good simulation of stratocumulus cloud (Martin et al. 2000).

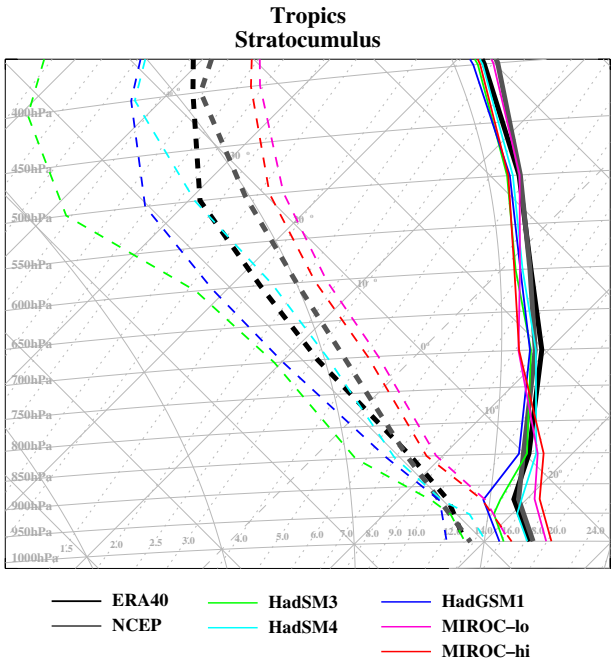

Fig. 5 Vertical profile of mean temperature (*solid*) and and specific humidity (*dashed*) for ERA-40 and NCAR/NCEP re-analyses, and for each GCM. Profiles are plotted on a standard UK tephigram for the tropical stratocumulus cloud regime

All of the models simulate regimes of deep convective cloud and thin cirrus. Several of the models simulate a deep convective tower and some cirrus, however there is little cloud with intermediate optical depths, which is characteristic of a convective anvil. With the exception of HadSM4 and HadGSM1, the GCMs tend to have CRF components which are weaker than observed for these regimes, whilst simulating the regime more frequently than is indicated by ISCCP.

All of the models have completely cloud-free grid-boxes more often than is observed by ISCCP, particularly ECHAM5 for which the lack of stratocumulus is largely substituted for with clear-sky. However, it is possible that some of the 'stratocumulus' observed over northern Africa is actually Saharan dust (e.g. Haywood et al. 2005). If this is the case, some of the discrepancy between the models and observations may be accounted for.

3.2 Ice-free extra-tropics

Over those spatio-temporal grid-points polewards of 20°N/S, which are free of ice or snow, five principal regimes may be identified from the observational clusters (Fig. 6). The observed regimes are: 'shallow cumulus'; 'stratocumulus'; 'cirrus'; 'mid-level cloud'; 'frontal' (although this regime will be referred to as 'frontal', the horizontal resolution of the data considered here are insufficient to explicitly resolve mesoscale frontal features).

The most frequently occurring cloud type is the shallow cumulus regime. Unlike the tropics, all of the models except ECHAM5 (which does not simulate this regime) produce some low cloud in the regime and the cluster centroid from HadSM4 may be considered reasonably realistic, given the satellite beam-filling caveats on the observations which were discussed above. As in the tro-

pics, most of the models simulate a stratocumulus regime with cloud which is thicker than observed, simulated less frequently than ISCCP in HadSM3 and more frequently than ISCCP in the MIROC models. Again, the simulated histograms from HadSM4 and HadGSM1 have the smallest Euclidean distance to those from ISCCP (ESM Table 2).

The observed regime with cloud top at mid-levels may be associated with cumulus congestus cloud capped by a stable layer at mid-levels and forming anvils (Webb et al. 2001). This normally occurs in cold air outbreaks behind cold fronts, as air from high latitudes is advected equatorwards and is heated from below by a relatively warm ocean surface. Examination of the geographical locations of the regime provides support for this interpretation as there is a maxima in RFO extending southeast from the Labrador Basin, south from the Bering Sea, and close to the mean Antarctic ice edge (Fig. 6). A contribution to this regime is also likely to occur from multi-layer cloud structures which exist ahead of frontal systems (Tselioudis and Jakob 2002). Only HadSM4 simulates a mid-level regime, and even in this model, the cloud top is lower than indicated by ISCCP. Simulating too little mid-level cloud is a general problem

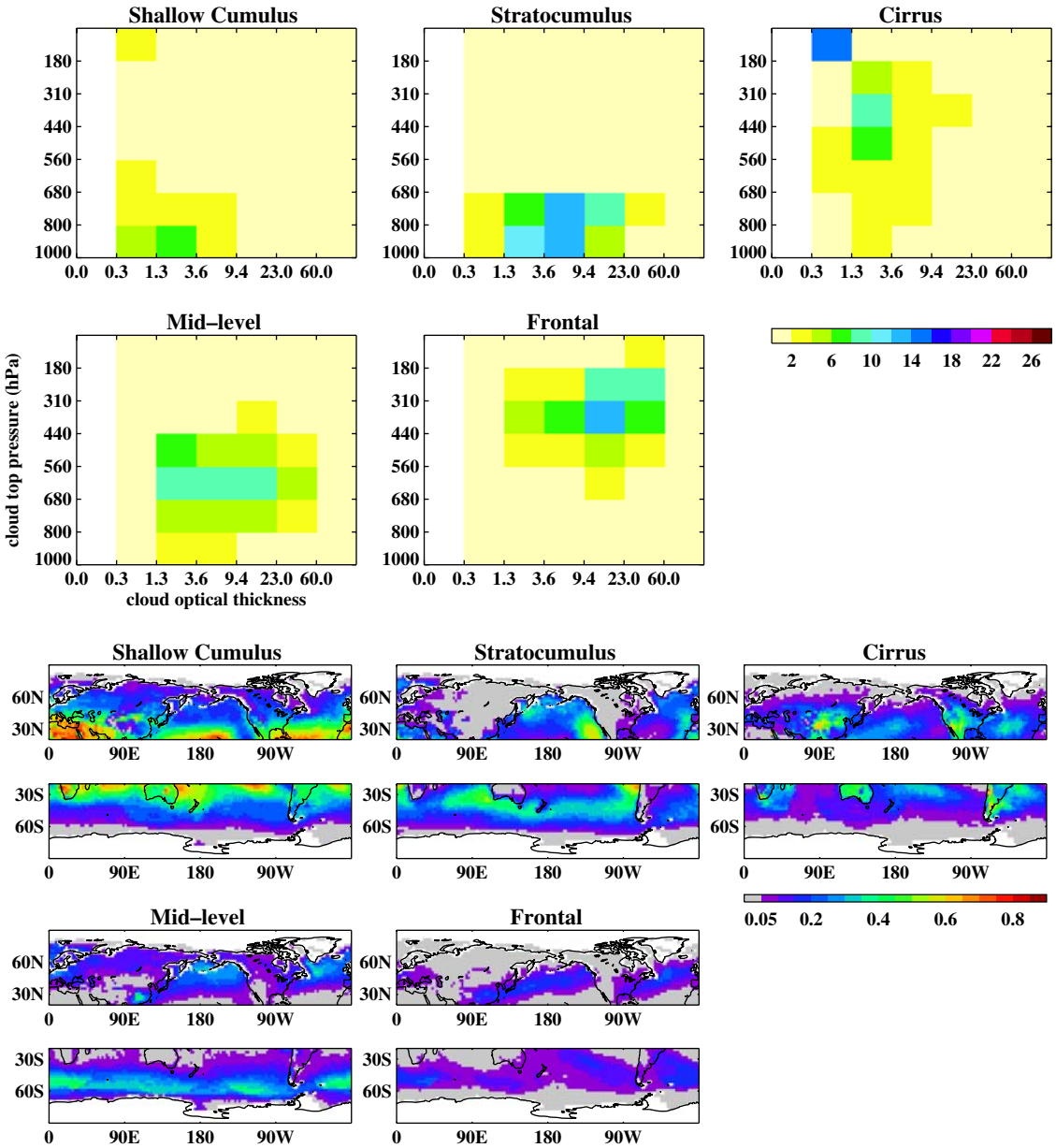

Fig. 6 *Top*: mean CTP-τ regime histograms resulting from clustering daily ISCCP data over the ice-free extra-tropics (polewards of 20°N/S). Some of the clusters have been combined to form a set of principal cloud regimes. *Bottom*: the temporal proportion of each grid-box (as a fraction of the total number of days) included in each regime

amongst contemporary GCMs (Webb et al. 2001; Zhang et al. 2005) and is likely to be mainly due to a lack of mid-level convective detrainment and poor simulation of multi-layer cloud structures (Tselioudis and Jakob 2002).

A regime of high-top, optically thick cloud associated with mid-latitude frontal systems is simulated by all of the GCMs although in several cases, the simulated regime contains cloud which is optically thicker than ISCCP (Fig. 7). The regime is simulated over six times more frequently in ECHAM5 than is observed by ISCCP with more than half of the spatio-temporal grid-points in the region being assigned as 'frontal' for this GCM. The high RFO of this regime in ECHAM5 results in the mid-latitude zonal-mean SCRF being too strong (Fig. 1).

Examination of the mean thermodynamic profiles for the frontal regime reveals considerable spread amongst the GCMs (Fig. 8), however ECHAM5 is unusual in that the profile is particularly dry throughout the troposphere. In addition, the re-analyses and most of the models suggest frontal cloud is associated with strong ascent, but EC-

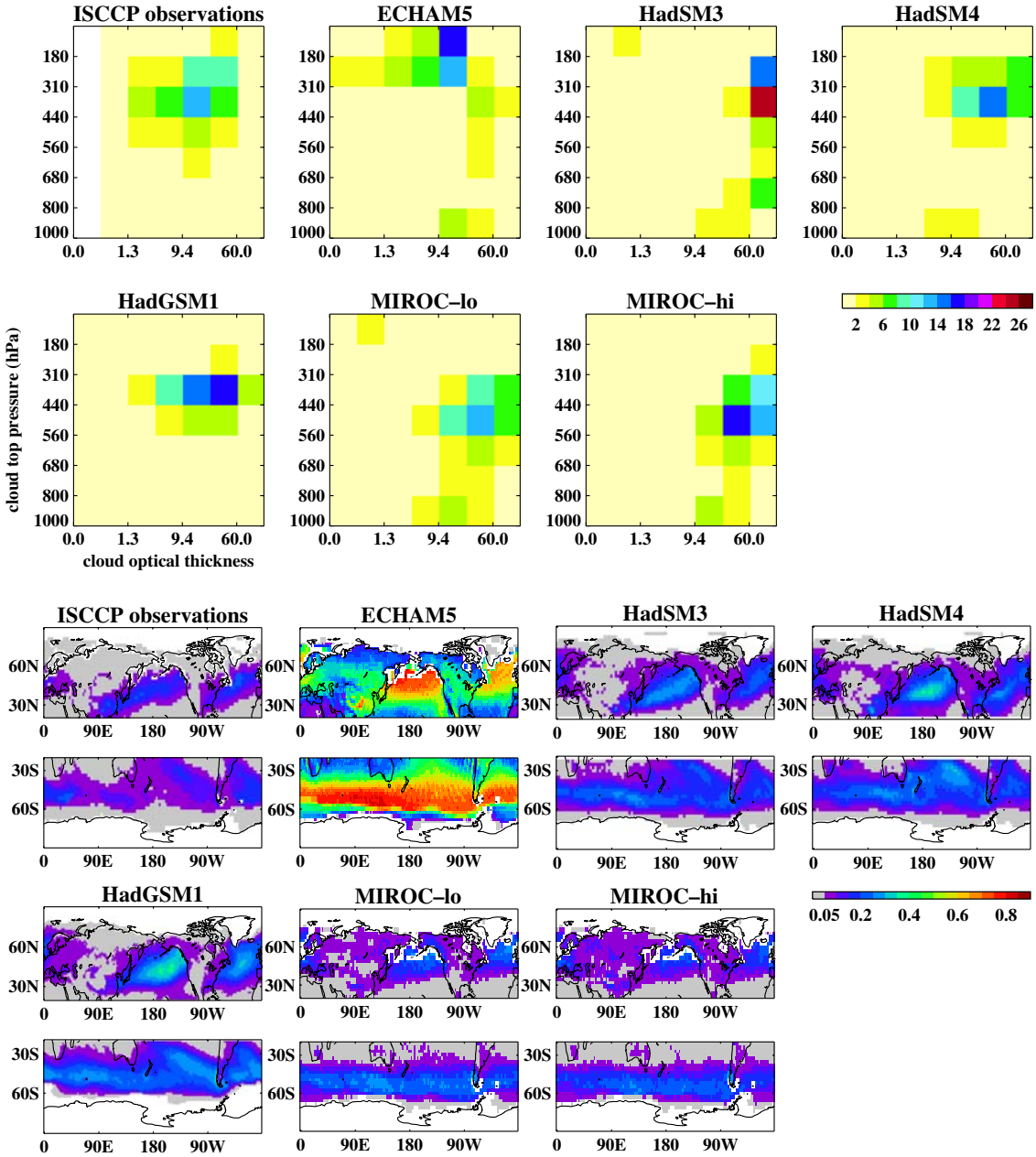

Fig. 7 *Top*: mean CTP-τ histograms for the extra-tropical frontal regime resulting from ISCCP and from each GCM. *Bottom*: the temporal proportion of each grid-box (as a fraction of the total number of days) included in the extra-tropical frontal regime

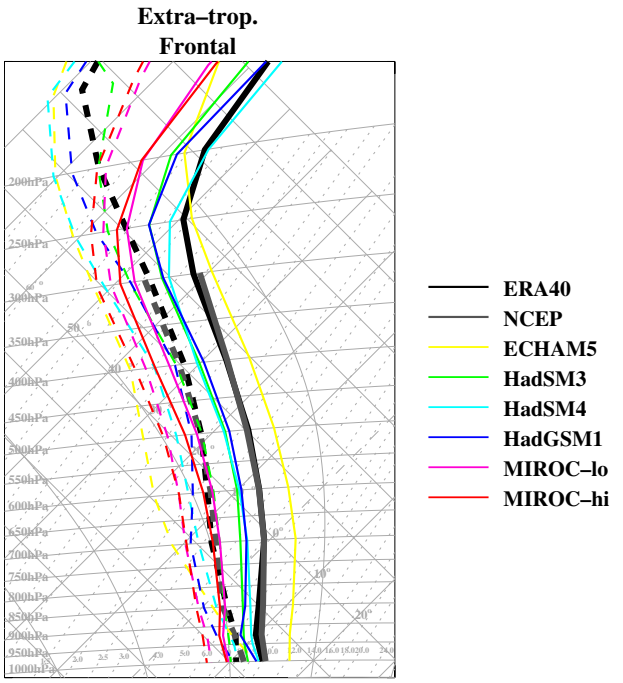

Fig. 8 As Fig. 5, but for the extra-tropical frontal regime

HAM5, on average, simulates weak descent for this regime (Table 4). Since these meteorological conditions would not normally be associated with frontal cloud, it is likely that frontal cloud is formed with ascent in moist conditions in ECHAM5, but persists (possibly enabled by having a prognostic cloud scheme) even when the air is subsequently undergoing descent. The persisting thick, high cloud would account for the considerable overestimate of the RFO of 'frontal cloud' in this model. The temperature profile for most of the models indicates a cold bias relative to the re-analyses, but ECHAM5 is warmer. This reflects differences in the geographical location of the regime with most of the models simulating the frontal cloud regime polewards of that observed by ISCCP over the Southern Ocean. In contrast, the widespread occurrence of thick, high cloud in ECHAM5 results in a warm bias for this model.

Table 4 Five hundred hectopascal of mean vertical velocity (hPa/day) (negative for ascent) for those grid-points assigned to the 'frontal' regime over the ice-free extra-tropics

ERA-40	–98.1
NCAR/NCEP	–84.2
ECHAM5	10.8
HadSM3	–97.0
HadSM4	–92.2
HadGSM1	–96.5
MIROC-lo	–82.5
MIROC-hi	–82.5

3.3 Ice and snow-covered regions

Those regions which are snow or ice-covered are now considered. The ISCCP observational datasets are less reliable over ice/snow at high latitudes, hence caution is required surrounding their use for evaluation in this region. Four principal cloud regimes may be identified from ISCCP and amongst the models: a low TCC regime; a regime of optically thicker cloud which is probably associated with decaying frontal systems, a regime of low cloud, and a regime dominated by cirrus (Fig. 9).

As in the ice-free extra-tropics, ECHAM5 simulates the 'frontal' regime more frequently and with a higher cloud top than ISCCP or any of the other GCMs. In common with ISCCP, most of the models simulate a regime with a low TCC occurring most frequently over over high ground (the East Antarctic ice sheet, Greenland and the Tibetan plateau), however ECHAM5 and MIROC-lo simulate the cirrus regime most frequently in these regions. All of the models except ECHAM5 have a regime of optically thicker low cloud which occurs most frequently over the Southern Ocean and Arctic Ocean, as indicated by ISCCP.

4 Climate change response

4.1 Climate change analysis methodology

Following Bony et al. (2004) and Williams et al. (2005), the average cloud radiative forcing over a region, $\overline{\mathrm{CRF}}$, can be obtained from clustered cloud regimes by:

$$\overline{\mathrm{CRF}} = \sum_{r=1}^{n\mathrm{regimes}} \mathrm{RFO}_r \mathrm{CRF}_r \tag{1}$$

where RFO_r is the RFO of regime r, and CRF_r is the average cloud radiative forcing in regime r. The change in cloud radiative forcing in response to doubling CO_2 is then given by:

$$\overline{\Delta\mathrm{CRF}} = \sum_{r=1}^{n\mathrm{regimes}} \mathrm{CRF}_r \Delta\mathrm{RFO}_r + \sum_{r=1}^{n\mathrm{regimes}} \mathrm{RFO}_r \Delta\mathrm{CRF}_r + \sum_{r=1}^{n\mathrm{regimes}} \Delta\mathrm{RFO}_r \Delta\mathrm{CRF}_r \tag{2}$$

The first term on the right-hand-side represents the contribution to $\overline{\Delta\mathrm{CRF}}$ from the change in the RFO of the regime; the second represents the change in CRF within the regime; and the third reflects a co-variation of RFO and CRF (this is found to be generally much smaller than the other two). In this section, the climate change response for each regime is decomposed into the components of Eq. 2.

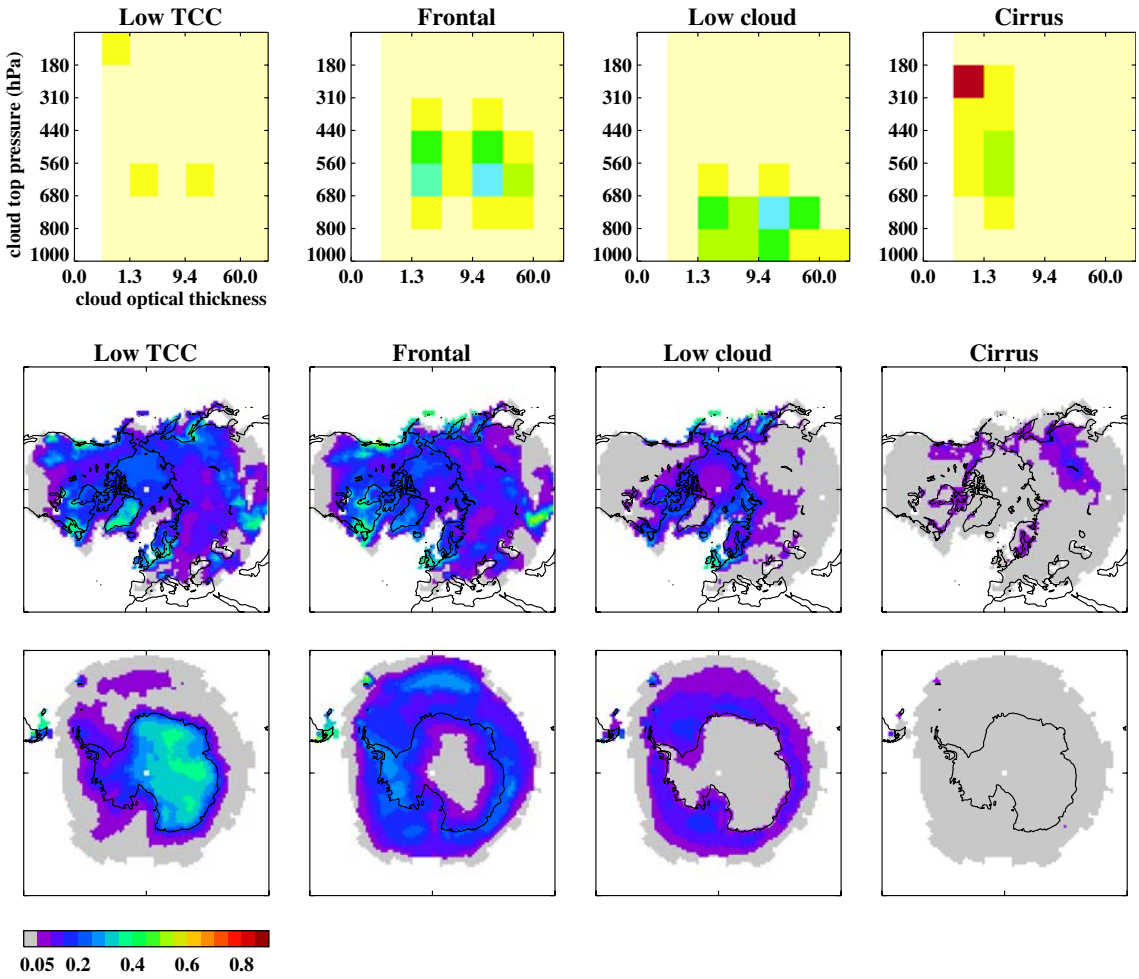

Fig. 9 *Top*: mean CTP-τ regime histograms resulting from clustering daily ISCCP data over spatio-temporal regions covered by snow/ice. Some of the clusters have been combined to form a set of principal cloud regimes. *Bottom*: the temporal proportion of each grid-box (as a fraction of the total number of days) included in each regime

The components are normalised by the global-mean climate sensitivity (the equilibrium surface temperature response to a doubling of CO_2) of the respective model in order to identify differences in local cloud feedback processes (Boer and Yu 2003; Webb et al. 2006).

Evaluation of the cloud response to climate change in the context of cloud regimes involves evaluation of each of the terms on the right-hand-side of Eq. 2. Clearly, only those regimes that are simulated by the model can be evaluated using this method. Evaluation of the present-day CRF_r and RFO_r in Eq. 2 is, in principle, straightforward and may be achieved through a comparison of the properties of the mean cloud regimes as simulated by the GCMs with observations, as undertaken in the previous section. Evaluation of ΔRFO_r and ΔCRF_r is more difficult. One approach may be to find proxies for the climate change response in present-day variability. It is intended that future work will address the evaluation of these change terms.

4.2 Tropics

The total response to doubling CO_2 over the tropical region is a positive $\overline{\Delta NCRF}$ in all of the models (Fig. 10). This appears to be primarily a result of a positive change in the NCRF of the cloud regimes, rather than a net change in the RFO of the regimes (i.e. the *black* bars in the left column of Fig. 10 are mainly formed from the *red* bar, with only a small contribution from the *blue* bar). Some of the GCMs do simulate changes in the RFO of the regimes, however the radiative response tends to approximately cancel on the large scale. For example, in HadSM3 there is a reduction in the RFO of cirrus, leading to a positive change in the $NCRF_r$ ΔRFO_r term (the SCRF component of this regime being slightly stronger than the LCRF), however the radiative effect is compensated for by an increase in the RFO of the shallow cumulus regime and stratocumulus.

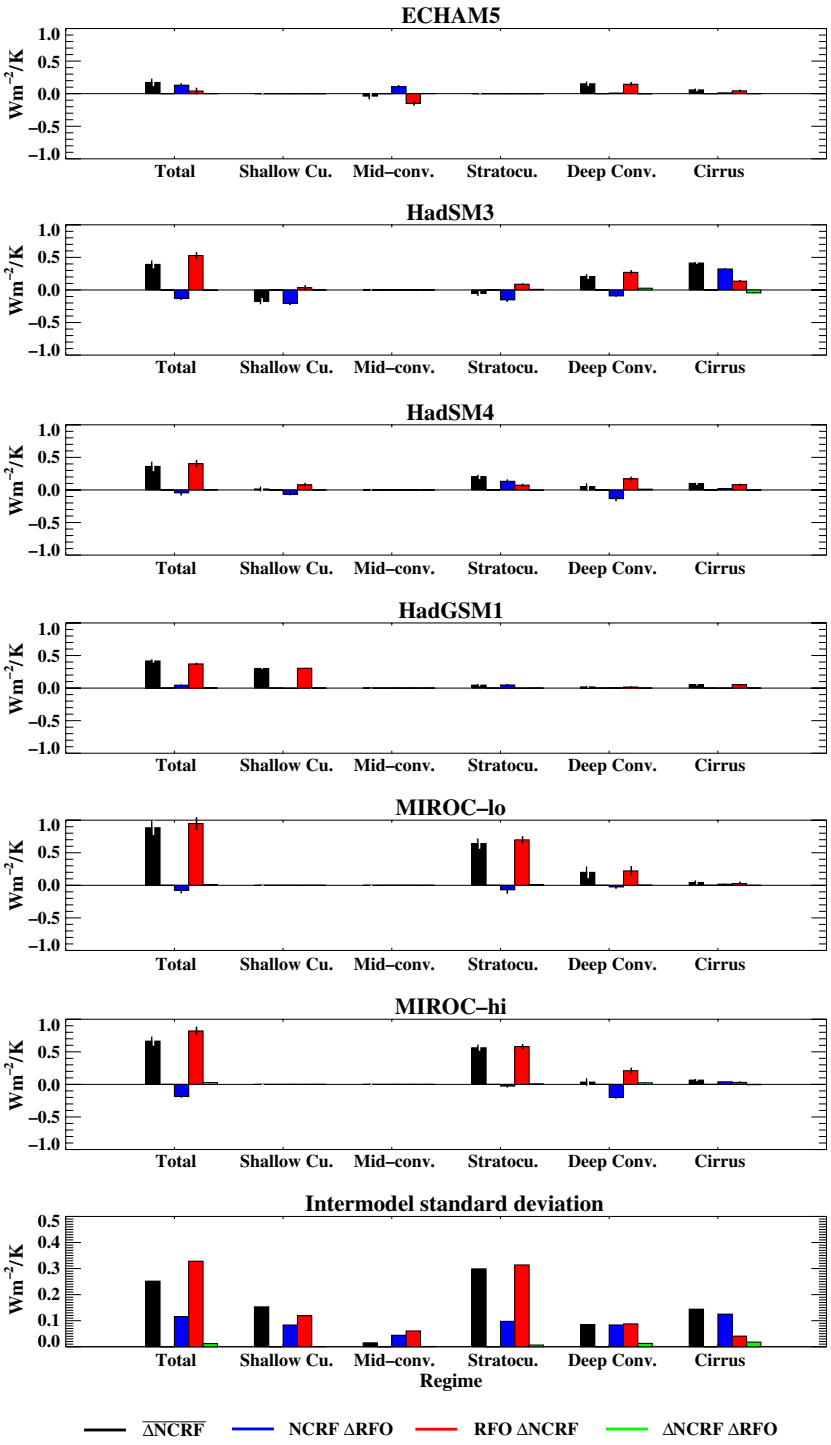

Fig. 10 The change in NCRF in response to doubling CO_2, normalised by the change in global mean surface temperature, decomposed into a contribution from each term in Eq. 2. Results are shown for each tropical cloud regime and for the total from all the regimes, hence the far left black bars are the $\overline{\Delta NCRF}$ for the region. The vertical lines at the end of each bar represent uncertainty estimates due to inter-annual variability

The inter-model variance in $\overline{\Delta NCRF}$ for the tropics can be seen to be dominated by differences in the response of stratocumulus, consistent with Webb et al. (2006) and Bony and Dufresne (2005). Again, this variance is largely due to differences in the radiative response of the regime rather than differences in the RFO response. The variance is mainly a result of the MIROC models having a large weakening of stratocumulus SCRF whilst the other GCMs show a much smaller weakening or no change. In HadSM3, HadSM4 and both MIROC models there is a reduction in the TCC of stratocumulus, with in-cloud τ and CTP of the regime being largely unchanged (i.e. when stratocumulus is formed in the $2 \times CO_2$ simulation, on average a smaller fraction of the grid-box is covered, or it is present for a shorter time in the day, than in the control). The regime-mean reduction in TCC (per K of warming) of all four

models is reasonably similar, however the radiative effect of this reduction in TCC is larger in HadSM3 and the MIROC models than HadSM4 because the optical depth of the regime is greater (Fig. 3). However, the overall radiative impact of the regime across the tropics (RFO_r $\Delta NCRF_r$) is considerably higher in the MIROC models (Fig. 10) due to the RFO of stratocumulus being much higher than in HadSM3 (Fig. 4). Therefore, much of the difference in the radiative response of stratocumulus to doubled CO_2 between models in this ensemble is due to differences in the present-day regimes simulated by the models, rather than differences in the change in grid-box cloud cover or cloud properties.

Given the evaluation of the control climate against ISCCP data presented in Sect. 3, the net radiative response of stratocumulus in the MIROC models appears excessive due to the RFO and optical depth of the regime being higher than observed. The sensitivity of the stratocumulus response to the regime-mean characteristics highlighted above, and its importance in contributing to variability in the net tropical response, suggests that future improvements in the simulation of this regime in GCMs may reduce the variation in the cloud feedback over the region 20°N–20°S.

4.3 Ice-free extra-tropics

As in the tropics, most of the variance in the $\overline{\Delta NCRF}$ for the extra-tropics arises from different changes in the radiative characteristics of the regimes. The variance in $\sum RFO_r\ \Delta NCRF_r$ largely arises from differences in the radiative response of the frontal cloud regime (Fig. 11). In particular, there is a strengthening of frontal cloud SCRF in ECHAM5 leading to a large negative $RFO_r\ \Delta NCRF_r$, whereas there is a small positive change in the other models. It is possible that a change in SCRF could occur through a simple latitudinal shift in a regime, due to the change in insolation. An additional bar has been added to Fig. 11 in which the SCRF is divided by the local insolation and multiplied by the regional annual-mean insolation before the difference between the simulations is calculated. This normalised change exhibits a similar level of inter-model variance, indicating that changes in the clouds, rather than in their location, drive the variance in RFO_r $\Delta NCRF_r$ between the models.

Changes in the TCC of the frontal regime are very small in all of the models, hence the radiative response is mainly a result of changes in τ and CTP. In all of the models, there is an increase in cloud top height associated with a lifting of the tropopause in a warmer climate (Wetherald and Manabe 1988) (Fig. 12). In all but the MIROC models, there is only a small $\Delta LCRF_r$ associated with this lifting as the difference in temperature between the cloud top and clear-sky average emission level in the lower troposphere is largely unchanged, i.e. the whole profile warms (Tompkins and Craig 1999). However, in the case of the MIROC models, the warming aloft is less than at lower levels, hence the temperature difference between the cloud top and clear-sky average emission level increases. This results in a considerable increase in LCRF (this is larger in MIROC-hi due to the higher climate sensitivity of this model - $\Delta LCRF_r$ normalised by the climate sensitivity is similar for the two models). The sign of $\Delta SCRF_r$ varies between the GCMs. In ECHAM5 and MIROC-lo there is a strengthening of SCRF due to an increase in the average optical depth of the regime. The very high RFO_r of frontal cloud in the ECHAM5 control (Fig. 7) results in the considerable negative RFO_r $\Delta NCRF_r$ response of this regime which can be seen in Fig. 11. The evaluation in Sect. 3 indicates that the high RFO of frontal cloud in ECHAM5 is in error, hence whilst the processes leading to a strengthening of the SCRF at grid-points containing frontal cloud may or may not be correct, the very large negative RFO_r $\Delta NCRF_r$ in ECHAM5 is questionable since it is a consequence of the excessive RFO.

The increase in LCRF described above in the MIROC models is partly offset in MIROC-lo by a strengthening of the SCRF associated with increased τ. In contrast, frontal cloud in MIROC-hi becomes optically thinner, with evidence of more cirrus being simulated within the frontal cloud regime. This leads to a weakening of SCRF in MIROC-hi, which adds to the strengthening LCRF, resulting

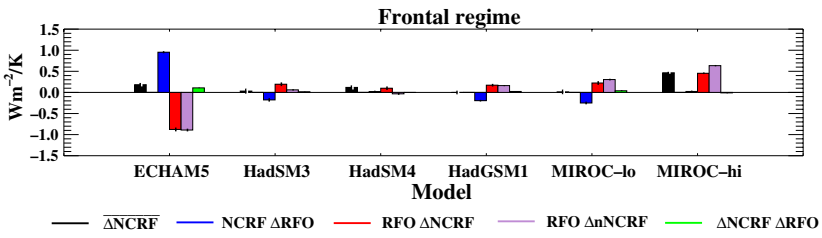

Fig. 11 The change in NCRF in response to doubling CO_2, normalised by the change in global mean surface temperature, for the frontal regime over the ice-free extra tropics. Also shown is *RFO* ΔnNCRF in which the SCRF component is normalised by the insolation (and multiplied by the mean insolation for the region). Note: the scale differs from Fig. 10

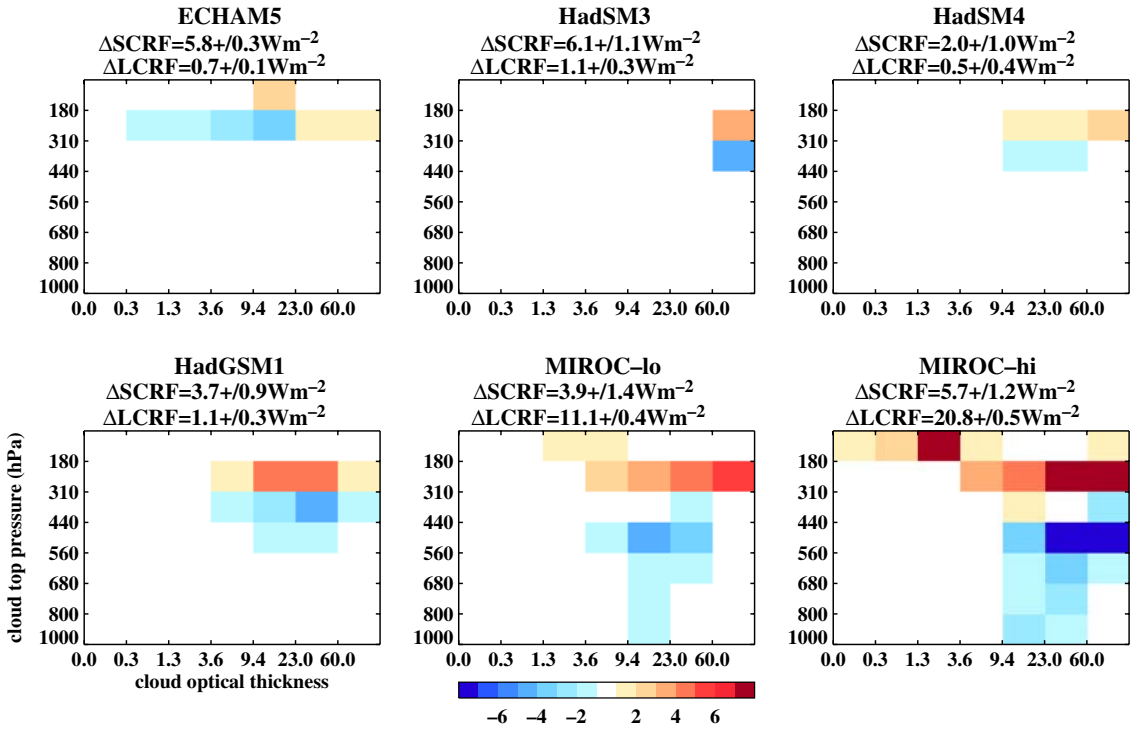

Fig. 12 Change in response to doubling CO_2 in regime-mean CTP-τ histograms for the frontal cloud regime over the ice-free extra-tropics. Also shown at the *top of each panel* is the mean $\Delta SCRF_r$ and $\Delta LCRF_r$

in a larger positive $\Delta NCRF_r$, contributing to the higher sensitivity of this model. This is consistent with the findings of Ogura et al. (2007), who attribute the cause of the different sensitivity between the MIROC models to an increase in cloud water in MIROC-lo compared with a reduction in MIROC-hi (due to enhanced auto-conversion and ice sedimentation in the latter) over the Southern Hemisphere mid-latitudes.

In ECHAM5 there is also a reduction in the RFO of frontal cloud leading to a positive $NCRF_r$ ΔRFO_r, however this is largely compensated for by an increase in the RFO of stratocumulus, much of which is geographically co-located (not shown). Therefore in ECHAM5, some frontal cloud is being replaced with stratocumulus which has a similar NCRF.

4.4 Ice and snow-covered regions

This region is analysed separately as the cloud masking of snow/sea-ice which reduces under climate warming can lead to a strengthening of SCRF even with no change in cloud (e.g. Cess et al. 1990; Winton 2005). Hence quantitative changes in CRF should be treated cautiously, however use of the ISCCP simulator data permits at least a qualitative assessment of whether the change in CRF is approximately consistent with any cloud change.

The $\overline{\Delta NCRF}$ over ice/snow-covered regions is large and negative (through a strengthening of SCRF) in all of the models (not shown). The overall response is again dominated by changes in the radiative properties of the regimes, but the regimes with the largest contribution to $\overline{\Delta NCRF}$ vary between the GCMs. In most of the models RFO_r $\Delta NCRF_r$ is largest in the low TCC and low cloud regimes. Whilst some of this response may be a cloud masking effect, analysis of changes in the regime histograms reveals that the cloud is tending to become optically thicker in the $2 \times CO_2$ climate (Fig. 13). This would suggest that much of the increase in SCRF is cloud related, rather than being due to clear-sky processes.

5 Potential to constrain the range of climate sensitivity

This section relates the response of each of the regimes computed in Sect. 4 with the global climate sensitivity to a doubling of $2 \times CO_2$. The analysis thus far has considered three geographical regions, however each region is a different size and the analysis presented has only considered sun-lit points. The total change in the CRF components from clustering each region are now multiplied by the fractional area of the globe covered by the region. In addition, the SCRF components for the polar region are

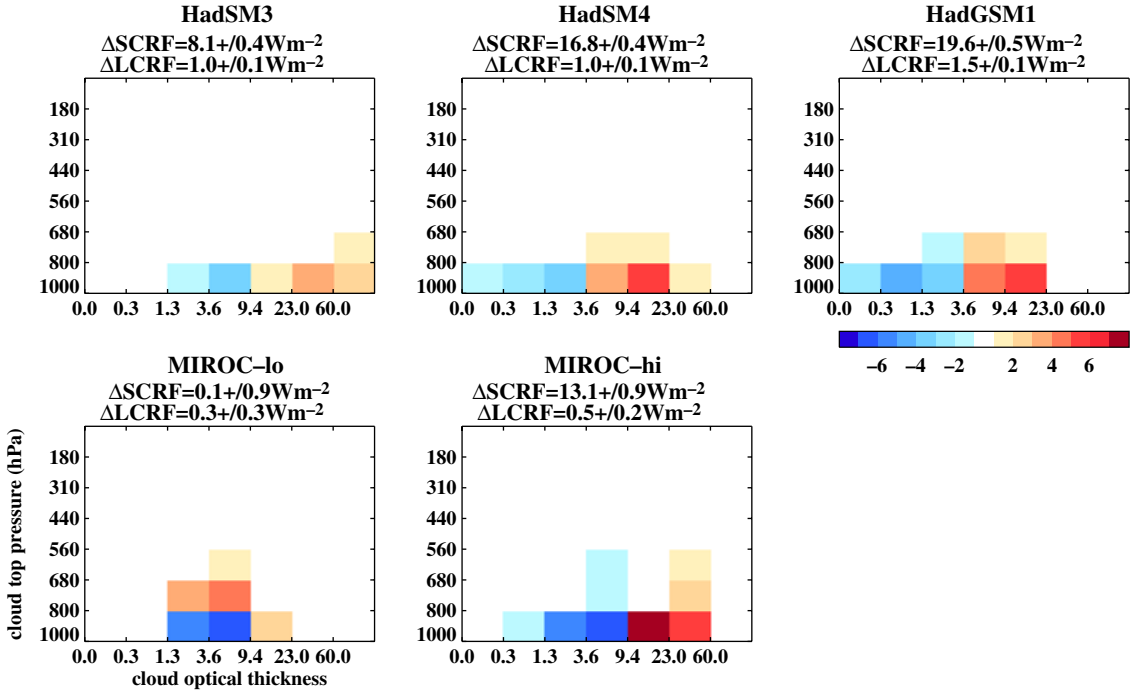

Fig. 13 Change in response to doubling CO_2 in regime-mean CTP-τ histograms for the low cloud regime over the region covered by snow/ice in the control. Also shown at the *top of each panel* is the mean $\Delta SCRF_r$ and $\Delta LCRF_r$. (ECHAM5 does not simulate this regime)

also multiplied by the spatio-temporal fraction of the region which is day-lit (i.e. the fraction for which ISCCP simulator diagnostics are produced). It is assumed that the mean LCRF obtained during the polar night is not very different to that obtained when it is light (this has been tested in several of the models and been found to be approximately true). The sum of the three regions therefore generates the global mean change in NCRF (Fig. 14).

With the exception of ECHAM5, all of the GCMs have a positive global $\overline{\Delta NCRF}$ which arises from consistently positive contributions from the tropics and extra-tropics and negative contribution of the ice covered regions. The largest contribution to the global mean variance in $\overline{\Delta NCRF}$ is from the ice-free extra-tropics and it has been shown above that a large contribution to the variance in this region results from differences in the response of frontal cloud. A secondary contribution to the global variance occurs from the tropical region, mostly resulting from differences in the stratocumulus response. The significant contribution to the variation in the global response from frontal cloud appears to contradict the result of Webb et al. (2006), whose global analysis suggests the main uncertainty arises from the response of low clouds. However, a large contribution to the variance found in the ensemble considered here is a result of the negative $\overline{\Delta NCRF}$ in ECHAM5 and this model is not analysed by Webb et al. (2006). The more general result of several studies (e.g. Webb et al. 2006; Bony and Dufresne 2005; Williams et al. 2006) is that most of the variation in the cloud response between models is in $\Delta SCRF$ rather than $\Delta LCRF$ and this remains valid for the ensemble considered here.

It has been shown qualitatively that much of the variance in the global mean change in NCRF between the GCMs is likely to be due to differences in the mean present-day regime characteristics, especially the regime RFO. We now aim to quantify the effect of these mean present-day differences on the variance of the cloud response, and on the range of climate sensitivity. Each of the terms in Eq. 2 is re-calculated with the present-day RFO_r and $NCRF_r$ parts replaced with the observed values for each regime as obtained from ISCCP. The $\Delta NCRF$ and ΔRFO parts are taken from the respective model as before. This provides an indication of the likely $\overline{\Delta NCRF}$ if the GCMs were able to simulate the mean present-day regime properties as observed, and assuming the response of each regime is as before.

The pale coloured bars in Fig. 14 show the resulting change for each component in Eq. 2, over each region and for the global-mean, using this 'observationally constrained' method. When constrained in this manner, *all* of the GCMs produce positive tropical and extra-tropical contributions to the global mean $\overline{\Delta NCRF}$ and negative contributions from the ice-covered regions. The standard deviation of the global-mean $\overline{\Delta NCRF}$ is just over a third of that obtained when using data from the GCMs. This mainly results from a reduction in the variance of the $\sum RFO_r$

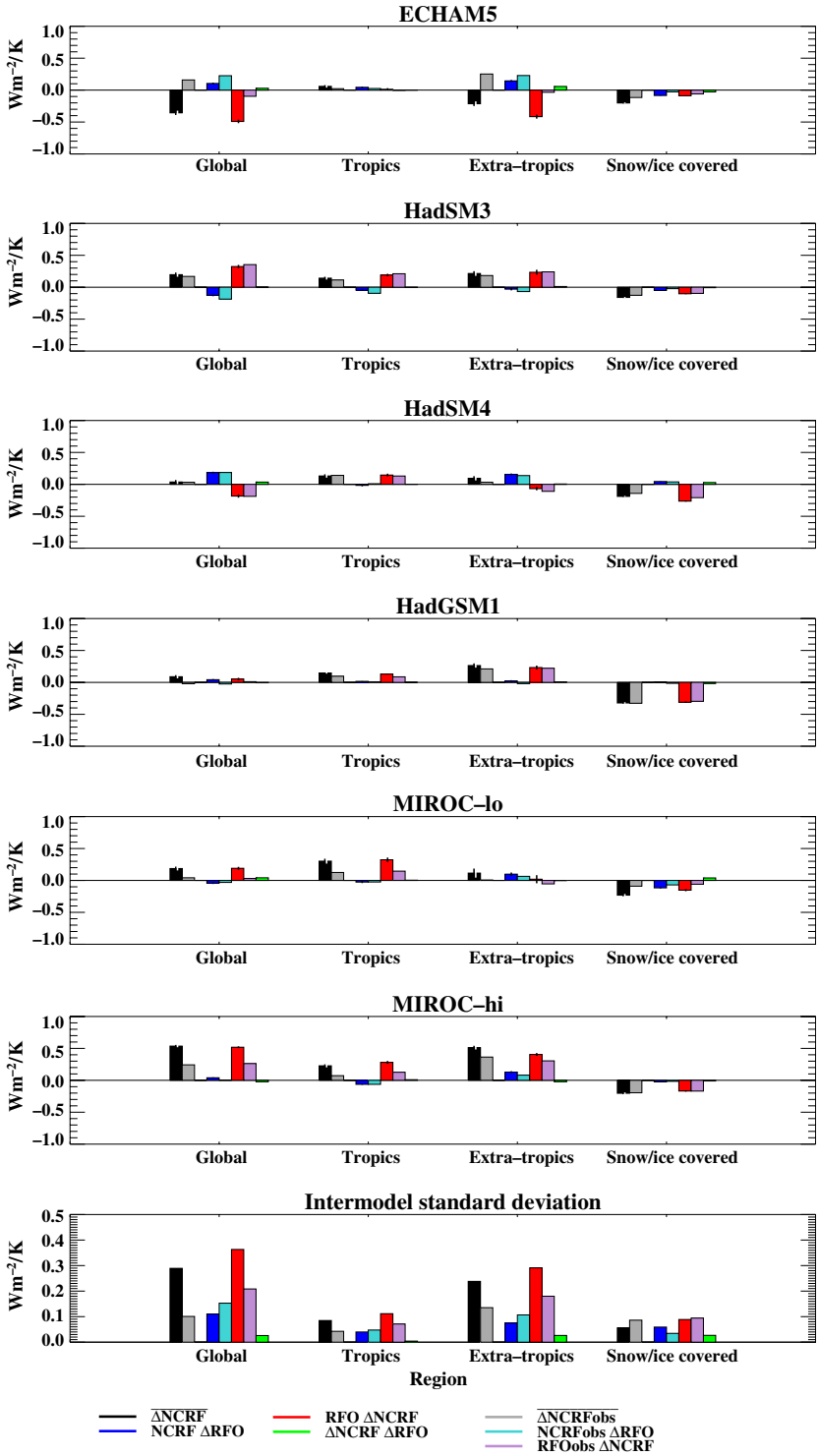

Fig. 14 Contribution of each region to the global-mean $\overline{\Delta NCRF}$, normalised by the change in global mean surface temperature response. The means for each region are weighted by the area of that region and the shortwave component is also weighted by the number of day-lit data-points. Also shown in the pale colours, are the changes in NCRF if observations of present-day RFO_r and $NCRF_r$ are used with the models' $\Delta NCRF_r$ and ΔRFO_r in the calculation of each term in Eq. 2

$\Delta NCRF_r$ term in Eq. 2. Whilst the impact of using the observed $NCRF_r$ in the $\sum NCRF_r \Delta RFO_r$ term is smaller, it does slightly increase the variance of the extra-tropics.

The difference in global-mean $\overline{\Delta NCRF}$ between the calculation constrained with observed data, and that obtained using just the model data, can be used to estimate the likely effect on the climate sensitivity parameter (λ), using the forcing due to doubling CO_2 from the respective GCM (e.g. Cess et al. 1990). Hence, the effect on the climate sensitivity can be estimated (Table 5). For the GCM ensemble analysed here, the effect of using observed present-day regime characteristics in the calculation of

Table 5 For each model: the difference in global-mean $\overline{\Delta NCRF}$ between that calculated for the model and that calculated when the observed present-day RFO_r and $NCRF_r$ is used in Eq. 2 (i.e. the *grey* bar minus the *black* bar in Fig. 14); the climate sensitivity parameter (λ) from the GCMs and that constrained by the observational data; the climate sensitivity (calculated as the difference in global-mean surface temperature between the 5-year periods of the $2 \times CO_2$ and control simulations analysed here—these figures may be slightly different from those reported elsewhere due to different periods of the simulation being averaged) and that constrained by the observational data

Model	Difference in $\overline{\Delta NCRF}$ (Wm^{-2}/K)	Model λ (Wm^{-2}/K)	Obs. constr. λ (Wm^{-2}/K)	Model clim. sens. (K)	Obs. constr. clim. sens. (K)
ECHAM5	0.49	1.21	0.72	3.3	5.6
HadSM3	0.17	1.06	0.89	3.5	4.2
HadSM4	0.03	1.00	0.97	3.7	3.8
HadGSM1	–0.11	0.83	0.94	4.6	4.1
MIROC-lo	–0.12	0.79	0.91	3.9	3.4
MIROC-hi	–0.19	0.48	0.67	6.5	4.7
Range		0.73	0.30	3.2	2.2
Std. dev.		0.25	0.12	1.2	0.8

$\overline{\Delta NCRF}$ is to increase the sensitivity of the models with the lowest climate sensitivity, and reduce the sensitivity of the higher sensitivity models. ECHAM5 goes from being the lowest sensitivity model to the highest, although the $\overline{\Delta NCRF}$ in the constrained case is not excessive compared with the other models. This implies that ECHAM5 has a relatively high clear-sky sensitivity which is partly offsetting the negative global mean ΔNCRF (Fig. 14), to give the model a sensitivity of 3.3 K.

Both the range and standard deviation of the climate sensitivities are reduced using the observational data compared with the actual GCM values. The range of climate sensitivity is found to reduced by almost a third, from 3.2 to 2.2 K. This suggests that if the present-day characteristics of the simulated cloud regimes can be made more realistic, the range of climate sensitivity amongst GCMs may be reduced. Of course, the calculation constrained by observations presented here is only illustrative. If the parametrizations in the GCMs are modified to provide simulated cloud regimes which are closer to observations, aspects of the cloud response to doubling CO_2 are also likely to change. In addition, only six GCMs are considered here. The diagnostic technique presented in this study needs to be applied to more models before it can be demonstrated that evaluation of the present-day cloud regimes in GCMs could be generally used to constrain the variation in climate sensitivity amongst models. However, the results presented above are encouraging.

6 Conclusions

A comparison exercise has been carried out on cloud regimes simulated by an ensemble of GCMs. A cloud clustering technique has been applied to consistently-produced model diagnostics to define principle cloud regimes globally, which can then be evaluated against observational data. Such an approach yields more detail about particular cloud types in error, than the more traditional approach of analysing the cloud in specific geographical regions (or zonal-means), where individual cloud types may be mixed temporally.

It is important for the analysis of the cloud response to climate change that at least some members of the model ensemble simulate each regime well. It is therefore of concern that none of the GCMs have a good representation of tropical trade cumulus in the 'shallow cumulus' regime. The simulation of cloud regimes with cloud tops at mid-levels also appears to be problematic for GCMs. Several of the models, and particularly HadSM3, simulate deep convective, stratocumulus and frontal cloud which is too optically thick. In the tropics, the MIROC models simulate stratocumulus more frequently than is observed, whilst HadSM3 only simulates this regime with a third of the frequency of ISCCP. In the extra-tropics, ECHAM5 simulates frontal cloud much more often than is observed by ISCCP. This appears to be due to high, optically thick cloud persisting in this model after meteorological conditions associated with frontal cloud have ceased. Overall, the mean cloud regime histograms and regime properties are closest to those observed in HadSM4 and HadGSM1.

The cloud response to climate change in the context of these principal cloud regimes has also been investigated. Whilst there are changes in the RFO of the regimes under climate change, the net radiative effect is found to be relatively small. It is not obvious that this would be the case since it is plausible that, for example, there could be an increase in the RFO of a high NCRF regime at the expense of one with a low NCRF (or of opposite sign). However in the GCMs investigated here, the product of NCRF and

ΔRFO largely cancel at global scales. Instead, most of the global cloud response to climate change arises through changes in the radiative properties of the individual regimes.

A large proportion of the variance in the global cloud response of the six models considered arises from differences in the radiative response of frontal cloud in the ice-free extra-tropics and from stratocumulus in the tropics. In both cases, much of the variation in the radiative response in this model ensemble is due to differences in the present-day regimes simulated by the models (specifically the RFO), rather than differences in the change in grid-box cloud cover or cloud properties. Since the evaluation of the present-day regimes indicates that the RFO of both frontal cloud in ECHAM5 and stratocumulus in the MIROC models is excessive when compared with ISCCP, the large response of these regimes in the three models may be questioned. This result suggests that work to improve the simulation of regime RFO in GCMs may be considered a priority.

Observations of the present-day RFO_r and CRF_r for the cloud regimes have been used with the model estimates of ΔCRF_r and ΔRFO_r to constrain the global-mean $\overline{\Delta NCRF}$, and climate sensitivity. For the GCM ensemble used in this study, use of observational data achieves a reduction in the range of climate sensitivity of almost a third. Therefore, the evaluation presented here shows potential to reduce the range of climate sensitivity amongst GCMs through evaluation and subsequent model improvement.

It is useful when evaluating GCMs to have a set of simple quantitative metrics by which to assess the models. The difference in global-mean $\overline{\Delta NCRF}$ between the GCM and that constrained by observations (left column of Table 5) provides one possible metric for evaluating the present-day cloud regimes. It implicitly weights the importance of the radiative climate change response of the regimes and, when decomposed, it highlights to model developers the regimes which require particular attention for reducing the current variation in cloud radiative response amongst GCMs. However, this metric does not include an assessment of whether the model actually simulates a particular regime and it does not evaluate the ΔRFO_r and ΔCRF_r terms in Eq. 2. Comparison of other simulated meteorological variables with observations/reanalyses, such as the regime mean temperature and humidity profiles will assist in assessing the realism of the regime.

This study has considered an ensemble of six GCMs in order to test the methodology and its potential for reducing uncertainty in cloud response to climate change. However, this ensemble is relatively small in comparison with the number of GCMs considered by the IPCC, and some of the models analysed here are reasonably similar to each other (e.g. the two MIROC models). It is therefore highly desirable that this methodology be applied to other GCMs. In order to achieve this, we continue to encourage modelling centres to submit daily data to the CFMIP and would welcome the CFMIP diagnostic lists becoming a standard part of the data requirements for the IPCC Fifth Assessment Report.

Acknowledgements This work was funded under the UK Government Meteorological Research programme. We thank Tomoo Ogura for submitting MIROC data and Johannas Quaas for submitting ECHAM5 data to CFMIP. Thanks go to Mark Ringer, Mark Webb, Catherine Senior, Tony Slingo, Christian Jakob, Michel Crucifix and William Ingram for useful discussions during this study and for their comments on early drafts of the paper. ISCCP data were obtained from the NASA Langley Research Center Atmospheric Sciences Data Center. NCEP reanalysis data were provided by the NOAA-CIRES Climate Diagnostics Center, Boulder, Colorado, USA, from their web site at http://www.cdc.noaa.gov.. ERA-40 data were obtained from ECMWF.

References

Anderberg MR (1973) Cluster analysis for applications. Academic, New York, p 359

Boer GJ, Yu B (2003) Climate sensitivity and response. Clim Dyn 20:415–429

Bony S, Dufresne JL (2005) Marine boundary layer clouds at the heart of cloud feedback uncertainties in climate models. Geophys Res Lett 32(20):L20806

Bony S, Dufresne JL, Le Treut H, Morcrette JJ, Senior CA (2004) On dynamic and thermodynamic components of cloud changes. Clim Dyn 22:71–86

Cess RD, Potter GL, Blanchet JP, Boer GJ, Delgenio AD, Dequé M, Dymnikov V, Galin V, Gates WL, Ghan SJ, Kiehl JT, Lacis AA, Le Treut H, Li ZX, Liang XZ, McAveney BJ, Meleshko VP, Mitchell JFB, Morcrette JJ, Randall DA, Rikus L, Roeckner E, Royer JF, Shlese U, Sheinin DA, Slingo A, Sokolov AP, Taylor KE, Washington WM, Wetherald RT, Yagai I, Zhang MH (1990) Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models. J Geophys Res 95:16601–16615

Cubasch U, Meehl GA, Boer GJ, Stouffer RJ, Dix M, Noda A, Senior CA, Raper SCB, Yap KS (2001) Projections of future climate change. In: Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden P, Dai X, Maskell K, Johnson CI (eds) Climate change 2001: the scientific basis. contribution of working group I to the third assessment report of the intergovernmental panel on climate change. Cambridge University Press, London, pp 525–582

Haywood JM, Allan RP, Culverwell I, Slingo A, Milton S, Edwards JM, Clerbaux N (2005) Can desert dust explain the outgoing longwave radiation anomaly over the Sahara during July 2003? J Geophys Res 110(D05105)

Jakob C, Tselioudis G (2003) Objective identification of cloud regimes in the Tropical Western Pacific. Geophys Res Lett 30(21)

Jakob C, Tselioudis G, Hume T (2005) The radiative, cloud and thermodynamic properties of the major Tropical Western Pacific cloud regimes. J Clim 18:1203–1215

Johns TC, Durman CF, Banks HT, Roberts MJ, McLaren AJ, Ridley JK, Senior CA, Williams KD, Jones A, Rickard GJ, Cusack S,

Ingram WJ, Crucifix M, Sexton DMH, Joshi MM, Dong BW, Spencer H, Hill RSR, Gregory JM, Keen AB, Pardaens AK, Lowe JA, Bodas-Salcedo A, Stark S, Searl Y (2006) The new Hadley Centre climate model HadGEM1: evaluation of coupled simulations. J Clim 19(7):1327–1353

K-1 model developers (2004) K-1 coupled model (MIROC) description. K-1 technical report 1. In: Hasumi H, Emori S (eds) Center for Climate System Research. University of Tokyo

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D (1996) The NCEP/NCAR 40-year reanalysis project. Bull Am Meteorol Soc 77(3):437–471

Klein SA, Hartmann DL (1993) The seasonal cycle of low stratiform clouds. J Clim 6(8):1587–1606

Klein SA, Jakob C (1999) Validation and sensitivities of frontal clouds simulated by the ECMWF model. Mon Weather Rev 127(10):2514–2531

Klein SA, Hartmann DL, Norris JR (1995) On the relationships among low-cloud structure, sea-surface temperature and atmospheric circulation in the summertime northeast Pacific. J Clim 8(5):1140–1155

Martin GM, Bush MR, Brown AR, Lock AP, Smith RNB (2000) A new boundary layer mixing scheme. Part II: tests in climate and mesoscale models. Mon Weather Rev 128(9):3200–3217

Martin GM, Ringer MA, Pope VD, Jones A, Dearden C, Hinton TJ (2006) The physical properties of the atmosphere in the new Hadley Centre Global Environmental Model (HadGEM1). Part I: model description and global climatology. J Clim 19:1274–1301

McAvaney BJ, Le Treut H (2003) The cloud feedback intercomparison project: (CFMIP). In: CLIVAR Exchanges—supplementary contributions. 26 March

Norris JR, Weaver CP (2001) Improved techniques for evaluating GCM cloudiness applied to the NCAR CCM3. J Clim 14:2540–2550

Ogura T, Emori S, Tsushima Y, Abe-Ouchi A, Kimoto M (2007) Climate sensitivity of a general circulation model with different cloud modelling assumptions. J Meteorol Soc Jpn (submitted)

Pope VD, Gallani ML, Rowntree PR, Stratton RA (2000) The impact of new physical parametrizations in the Hadley Centre climate model—HadAM3. Clim Dyn 16:123–146

Ringer MA, Allan RP (2004) Evaluating climate model simulations of tropical cloud. Tellus 56:308–327

Ringer MA, McAvaney B, Andronova N, Buja L, Esch M, Ingram W, Li B, Quaas J, Roeckner E, Senior C, Soden B, Volodin E, Webb M, Williams K (2006) Global mean cloud feedbacks in idealized climate change experiments. Geophys Res Lett 33:L07718

Roeckner E, Bauml G, Bonaventura L, Brokopf R, Esch M, Giorgetta M, Hagemann S, Kirchner I, Kornblueh L, Manzini E, Rhodin A, Schlese U, Schulzweida U, Tompkins A (2003) The atmospheric general circulation model ECHAM 5. Part I: model description. Technical Report 349, Max Planck Institute for Meteorology

Rossow WB, Schiffer RA (1991) ISCCP cloud data products. Bull Am Meteorol Soc 72:2–20

Rossow WB, Walker AW, Beuschel DE, Roiter MD (1996) International Satellite Cloud Climatology Project (ISCCP) documentation of new cloud datasets. WMO/TD 737, World Meteorological Organisation, p 115

Rossow WB, Tselioudis G, Polak A, Jakob C (2005) Tropical climate described as a distribution of weather states indicated by distinct mesoscale cloud property mixtures. Geophys Res Lett 32(21)

Senior CA, Mitchell JFB (1993) Carbon dioxide and climate: The impact of cloud parameterization. J Clim 6:393–418

Soden BJ, Held IM (2006) An assessment of climate feedbacks in coupled ocean-atmosphere models. J Clim 19:3354–3360

Tompkins AM, Craig GC (1999) Sensitivity of tropical convection to sea surface temperature in the absence of large-scale flow. J Clim 12(2):462–476

Tselioudis G, Jakob C (2002) Evaluation of midlatitude cloud properties in a weather and a climate model: dependence on dynamic regime and spatial resolution. J Geophys Res 107(D24):4781

Tselioudis G, Zhang Y, Rossow WB (2000) Cloud and radiation variations associated with northern midlatitude low and high sea level pressure regimes. J Clim 13(2):312–327

Wang J, Rossow WB, Uttal T, Rozendaal M (1999) Variability of cloud vertical structure during ASTEX observed from a combination of rawinsonde, radar, ceilometer, and satellite. Mon Weather Rev 127(10):2484–2502

Webb M, Senior C, Bony S, Morcrette JJ (2001) Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models. Clim Dyn 17:905–922

Webb MJ, Senior CA, Sexton DMH, Ingram WJ, Williams KD, Ringer MA, McAvaney BJ, Colman R, Soden BJ, Gudgel R, Knutson T, Emori S, Ogura T, Tsushima Y, Andronova NG, Li B, Musat I, Bony S, Taylor KE (2006) On the contribution of local feedback mechanisms to the range of climate sensitivity in two GCM ensembles. Clim Dyn 27(1):17–38

Wetherald RT, Manabe S (1988) Cloud feedback processes in a general circulation model. J Atmos Sci 45:1397–1415

Williams KD, Senior CA, Mitchell JFB (2001) Transient climate change in the Hadley Centre models: the role of physical processes. J Clim 14(12):2659–2674

Williams KD, Ringer MA, Senior CA (2003) Evaluating the cloud response to climate change and current climate variability. Clim Dyn 20:705–721

Williams KD, Senior CA, Slingo A, Mitchell JFB (2005) Towards evaluating cloud response to climate change using clustering technique identification of cloud regimes. Clim Dyn 24:701–719

Williams KD, Ringer MA, Senior CA, Webb MJ, McAvaney BJ, Andronova N, Bony S, Dufresne JL, Emori S, Gudgel R, Knutson T, Li B, Lo K, Musat I, Wegner J, Slingo A, Mitchell JFB (2006) Evaluation of a component of the cloud response to climate change in an intercomparison of climate models. Clim Dyn 26:145–165

Winton M (2005) Simple optical models for diagnosis surface – atmosphere shortwave interactions. J Clim 18:3796–3805

Zhang Y, Rossow WB, Lacis AA, Oinas V, Mishchenko MI (2004) Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: refinements of the radiative transfer model and input data. J Geophys Res 109(D19,105)

Zhang MH, Lin WY, Klein SA, Bacmeister JT, Bony S, Cederwall RT, Del Genio AD, Hack JJ, Loeb NG, Lohmann U, Minnis P, Musat I, Pincus R, Stier P, Suarez MJ, Webb MJ, Wu JB (2005) Comparing clouds and their seasonal variations in 10 atmospheric general circulation models with satellite measurements. J Geophys Res 110(D15)